Final Report for Period: 10/2009 - 09/2010
Submitted on: 12/14/2010
Principal Investigator: Koltchinskii, Vladimir .
Award ID: 0624841
Organization: GA Tech Res Corp - GIT
Submitted By:
Koltchinskii, Vladimir - Principal Investigator
Title:
MSPA-MCS: Sparsity in High-Dimensional Learning Problems

Project Participants

Senior Personnel

Name: Koltchinskii, Vladimir
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Yuan, Ming
Worked for more than 160 Hours: Yes
Contribution to Project:

Post-doc

Graduate Student

Name: Ruan, Lingyan
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Minsker, Stas
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Ni, Kai
Worked for more than 160 Hours: Yes
Contribution to Project:

Undergraduate Student

Technician, Programmer

Other Participant

Research Experience for Undergraduates

Organizational Partners

Other Collaborators or Contacts

Philippe Rigollet, Post Doc of the PI, School of Mathematics, Georgia Tech (currently, Assistant Professor at Princeton)

Evarist Gine, Professor, Departments of Mathematics and Statistics,
University of Connecticut

Alexandre Tsybakov, Professor, Laboratory of Probability and Stochastic Models, University
Paris VI and CREST, France

Shahar Mendelson, Professor, Department of Mathematics, Technion, Israel

Roshan Joseph, Associate Professor, School of Industrial and Systems Engineering, Georgia
Tech

Renato Monteiro, Professor, School of Industrial and Systems Engineering, Georgia Tech

Ali Eciki, PhD student, School of Industrial and Systems Engineering, Georgia Tech

Pedro Rangel, PhD student, School of Mathematics, Georgia Tech

Zhaosong Lu, Assistant Professor, Department of Statistics, Carneige Mellon University

Yi Lin, Department of Statistics, University of Wisconsin

Hui Zou, Department of Statistics, University of Minnesota

Marten Wegkamp, Department of Statistics, Florida State University

Tony Cai, Department of Statistics, Wharton School, University of
Pennsylvania

Activities and Findings

Research and Education Activities:

The main goal of this project is to develop a general mathematical framework needed to describe sparsity phenomena in high-dimensional problems of learning theory (such as regression and classification) and to study specific penalization techniques that provide a way to recover a sparse solution of the problem based on empirical data (given that such a solution exists).

The project started in October, 2006 and ended in September, 2010. The report summarizes the main activities over this period.

The PI and the co-PI have pursued several lines of research described below in some detail.

1. The PI has studied a number of problems related to sparse recovery in the case of finite dictionary using penalized empirical risk minimization with l_p-type penalties for the values of p close to 1 (see problems 1-4 in the proposal). Based on the results of this research, the PI has written a paper that appeared in the 'Annales de'l Institut H. Poincare' in 2009.

2. The PI has also studied the connections of the approach to sparse recovery described in the proposal with another method developed recently by Candes and Tao (the Dantzig selector).
The PI has written another paper based on this work. It

was published in 'Bernoulli' in 2009.

3. The PI and the co-PI have been working jointly on a circle of questions related to recovery of sparse solutions of learning problems in the case of finite 'dictionaries' that consist not of functions, but rather of Reproducing Kernel Hilbert Spaces (RKHS). It happened that the techniques developed by the PI in the case of simple dictionaries can be extended to this more general class of learning problems and that penalized empirical risk minimization with l_p-type penalties still provides adaptation to unknown sparsity of the problem. A brief description of these activities, primarily, their connection to regression and classification problems with prior information about grouping of covariates (features), was given in the proposal (see Problem 7). However, a more detailed analysis undertaken by the PI and the co-PI has shown that a number of interesting statistical problems such as sparse recovery in additive models and aggregation of kernel machines also fit this general framework. The first results in this direction were presented at the COLT 2008 conference in Helsinki and the paper on these results has appeared in the COLT 2008 volume.

The PI and the co-PI also developed another approach to the problem that is based on dual penalization with a combination of two l_1-type norms. One of them is based on the RKHS-norms of the components and it is used to enforce their smoothness. Another one is based on the empirical L_2-norms of the components and it is used to enforce their sparsity. In the case of classical additive models, such an approach was recently studied by Ravikumar, Liu, Lafferty and Wasserman and by Meier, van de Geer and Buhlmann ('sparse additive models'). In the context of multiple kernel models, the PI and the co-PI developed a version of such a method that is adaptive to the unknown smoothness of the components. Their paper on this subject has been published in the Annals of Statistics.

4. The PI and the co-PI have been working on several problems related to estimation of large covariance matrices and high-dimensional principal component analysis under various sparsity assumptions and are developing an approach to these problems based on the Dantzig selector and other recent methods of sparse recovery.

5. The PI was working (jointly with Evarist Gine, University of Connecticut) on generalization error bounds for kernel machines based on the approach to analysis of penalized empirical risk minimization developed by the PI in his paper that has been published in the 'Annales de'l Institut H. Poincare'.

6. The PI was working with his post doc Philippe Rigollet (currently, an associate professor at Princeton) on a circle of problems related to sparsity in ensemble learning (see Section 2.2 of the project description).

7. The PI has developed a method of sparse recovery of a function in a convex hull of a given dictionary based on negative entropy penalization. His paper on this subject was published in the Annals of Statistics in 2009.

8. The PI is working with his PhD student Stas Minsker (who has been supported by the grant) on sparse recovery

problems in the case of infinite dictionaries including
sparse mixtures recovery in density estimation and regression
framework (which is related to Section 2.2 of the proposal).
This is a topic of PhD thesis of Stas Minsker.
Jointly with Stas Minsker, the PI published a paper
on this subject in COLT 2010. A more detailed paper
on this circle of problems is currently in preparation.

9. The co-PI has studied the consistency issue for popular variable selection methods in linear regression including LASSO, LARS, elastic net and nonnegative garrote. In a paper (co-authored with Yi Lin) published in Journal of Royal Statistical Society, Series B, 2007, it was shown that nonnegative garrote is the only method that is consistent in variable selection and enjoys similar computational advantages as the LARS.

10. The co-PI also investigated new ways of exploring sparsity structure in multivariate linear regression where multiple responses are considered simultaneously. By imposing a penalty that is proportional to the coefficient matrix's Ky-Fan norm (or equivalently nuclear norm, trace norm), a method that shrinks the estimate to a factor model was proposed. The findings have been reported in papers recently appeared in Journal of Royal Statistical Society, Series B and Mathematical Programming. (co-authored with Renato Menteiro, Ali Ekici and Zhaosong Lu).

11. The co-PI has studied ways to accommodate marginal or heredity principles in variable selection. Both types of principles are widely used in practice to achieve meaningful model selection. However, most of the practical approaches are ad hoc in nature and little is known about their properties. In a paper to appear in Technometrics, the co-PI, together with Roshan Joseph and Yi Lin, proposed a computationally efficient strategy for this purpose. The paper has also been selected as the Technometrics invited paper in the 2007 INFORM annual meeting held in Seattle. More recently, the co-PI has also studied how more general constraints than heredity and marginal principles could be enforced in variable selection. Some of the findings have been included in a paper that appeared recently in the Annals of Applied Statistics.

12. During his visit to Technion in May, 2009, the PI started
a joint work with Shahar Mendelson on the development of normal
approximation of LASSO-type estimators in high dimensional
problems (known results of Knight and Fu are in the case of
large sample size and fixed dimension of the data).
Last spring, Evarist Gine (University of Connecticut)
has also joined this effort. Some results have been already
obtained, but still one has to overcome substantial difficulties
in order to show these results under the optimal assumptions
on the degree of sparsity. These difficulties are related to the fact that,
in the existing bounds on the accuracy of normal approximation
in the multivariate case, the dependence on the dimension is
not optimal (at least, in this specific application). The problem
is still under investigation.

13. The PI has been working on the development of excess risk
bounds in active learning, an approach to classification in which
learning algorithms can modify the design distribution in the
process of learning which leads to the reduction of the amount
of labeled data needed to achieve certain accuracy of learning.
The PI has published a paper on this subject in the Journal

of Machine Learning Research. Stas Minsker (a PhD student of the PI) is finishing his paper on computationally tractable versions of active learning algorithms and on lower bounds in active learning.

14. The PI has been working on low rank matrix estimation problems. One of the projects deals with estimation of density matrix in quantum state tomography. The approach is based on von Neumann entropy penalization. Another project (joint with Alexandre Tsybakov and Karim Lounici) is about more traditional matrix regression problems (including noisy matrix completion). The approach is based on nuclear norm penalization and the goal is to obtain sharp oracle inequalities (with the leading constant 1) for this method.

15. The PI and the co-PI have organized a seminar at Georgia Tech on Sparsity in High Dimensional Problems in Learning Theory.
The seminar was attended by several faculty members, post docs, visitors and graduate students from Mathematics, ISYE and Electrical Engineering. In particular, Justin Romberg (EE), Xiaoming Huo (ISYE), Philippe Rigollet (post doc of the PI, Mathematics), Soyeon Chun (graduate student from ISYE), Stas Minsker (PhD student of the PI) were among active participants of the seminar. The discussions at the seminar have led to joint research projects between its participants.

16. The Workshop on Sparsity in High Dimensional Statistics and Learning Theory took place in March, 2008. Justin Romberg and Xiaoming Huo joined the PI and the co-PI on the Organizing Committee. The workshop was supported by this NSF grant and also by Georgia Tech (College of Sciences, ISYE and ECE). The following topics were covered:

Asymptotic Geometric Analysis and High Dimensional Probability in Sparse Recovery Problems;

Model Selection and Oracle Inequalities for Statistical Models of Sparsity;

Penalization Methods, Dantzig Selector and Other Methods of Sparse Recovery;

Greedy Approximation Methods;

Concepts of Sparsity and Complexity in Learning Theory;

Sparsity and Generalization Bounds for Learning Machines.

The list of invited speakers included Andrew Barron (Yale), Sara van de Geer (ETH), Alexandere Tsybakov and Albert Cohen (Paris 6), Ron DeVore (South Carolina), Jianqing Fan (Princeton), Jon Wellner (University of Washington), Bin Yu (Berkeley) and several other researchers (both well known and young). Two PhD students were also among the speakers. The workshop was attended by more than 40 people (including many graduate students).

17. The PI and the co-PI have been attending a number of conferences and giving a number of invited talks related to their research supported by this grant, inculding the following:

- invited talk by the PI at the conference on Mathematical Statistics in Luminy, France, December, 2006;

- invited talk by the PI at the Workshop on Geometry, Random Matrices and Statistical Inference at SAMSI in January, 2007;

- invited talk by the PI at Joint Statistical Meeting in Salt Lake City, Utah in August, 2007;

- invited talk by the PI at the Workshop in Analysis and Probability, Texas A&M University, College Station, Texas in August, 2007;

- invited talk by the PI at the Workshop on Approximation and Learning in High Dimensions, Texas A&M University, Texas, in October, 2007;

- invited talk by the PI in the Distinguished Seminar Series, University of Genoa, Italy in October 2007;

- invited talks by the PI and the co-PI at the Joint Statistical Meeting in Salt Lake City, Utah, August, 2007;

- invited poster presentation by the co-PI at AISTAT 2007 in Puerto Rico, March, 2007.

- a talk by the PI and the co-PI at COLT 2008 in Helsinki, July, 2008

- invited talk by the PI at the Workshop on Stochastic Inequalities and their applications at Banff, Canada, June 2009.

- invited talk by the co-PI at the Workshop on Sparsity in Machine Learning and Statistics at Cumberland Lodge, UK, April 2009.

- invited talk by the co-PI at the 1st IMS Asian Pacific Rim Meeting, Seoul, South Korea, June 2009.

- invited talk by the co-PI at the 2nd IMS China Meeting, Weihai, China, June 2009.

- invited talk by the co-PI at the JSM, Washington DC, August 2009.

- invited talk by the PI at a Workshop on 'Modern nonparametric statistics: going beyond asymptotic minimax' at Oberwolfach, Germany in March, 2010.

- invited talk by the PI at a Workshop on Probability and Geometry in High Dimensions in Marne-la-Vallee, France in May, 2010.

- a talk by the PI at the 23rd Conference on Learning Theory COLT 2010 in Haifa, Israel in June 2010.

- invited talk by the PI at a conference 'From probability to statistics and back: high-dimensional models and processes' (a conference to honor Jon Wellner) in Seattle, July, 2010

- invited talks by the PI at Probability, Statistics and Computer Science seminars in University of Connecticut (March, 2007), University of Minnesota (April, 2007), University Paris VI and University of Lille, France (May, 2007), Uiversity of

California, Berkeley (April, 2008), Technion, Israel (May, 2009), Cambridge, UK (June, 2010), Yale (September, 2010), Princeton (September, 2010)

In May-June 2008, the PI and the co-PI participated in the program on Statistical Theory and Methods for Complex High Dimensional Data as visiting Fellows at Isaac Newton Institute for Mathematical Sciences, in Cambridge, England and gave invited presentations during their stay in Cambridge and at the closing workshop.

The PI has delivered a series of lectures on his research at Ecole d'ete de Probabilites Saint-Flour (July, 2008). The lecture notes of this well known School are traditionally published in a special series in Lecture Notes in Mathematics, Springer.

The PI also gave an invited 4 hours tutorial on Excess Risk Bounds in Machine Learning at the Summer School on Theory and Practice of Computational Learning in Chicago, June, 2009 (Organizers: Belkin, Niyogi and Smale).

Two graduate students have joined the team in August, 2007: Stas Minsker who is a PhD student of the PI and Linyan Ruan who is a PhD student of the co-PI. Both of them were supported in part by this grant. Two more students, Pedro Rangel and Kai Ni, have joined the team in 2009. Kai Ni was also supported in part by this grant.

Philippe Rigollet was a post doc of the PI (supported by Georgia Tech) in 2007-2008. He accepted a tenure-track faculty position at Princeton in June, 2008.

Findings:

1. The paper by the PI 'Sparsity in Penalized Empirical Risk Minimization' (which has appeared in the the 'Annales de'l Institut H. Poincare' in 2008) studies the problem of penalized empirical risk minimization (including regression and classification) over the linear span of a given dictionary consisting of a large number N of functions, the complexity penalty being the p-th power of the l_p-norm of the vector of coefficients times a regularization parameter. A number of probabilistic inequalities, showing that, when the solution of the 'true' risk minimization problem is 'sparse', the solution of its empirical version is, with a high probability, 'approximately sparse', have been proved in this paper. As a consequence of these 'sparsity bounds', the PI established oracle inequalities on the excess risk of the penalized empirical risk minimizers that show some form of optimality of the empirical solution and its adaptivity to unknown sparsity of the problem. The most interesting feature of these results is that for the values of p in the l_p-norm larger than 1, but close to 1 (within 1/log N from 1) it is possible (by exploiting properly strong convexity and smoothness of the penalty) to show sparsity bounds and oracle inequalities with almost no assumptions on the 'well posedness' of the dictionary. At the moment, it seems to be impossible to reproduce this type of results for other methods of sparse recovery such as LASSO (l_p penalization with p=1), the Dantzig selector of Candes and Tao, etc, where various rather strong 'well posedness' assumptions are needed (this is discussed in the paper by the PI 'The Dantzig selector and sparsity oracle inequalities' that appeared in 'Bernoulli' in 2009).

2. The approach developed by the PI in the case of 'simple dictionary'
(a set of N given functions) has been extended to a more general
framework in the joint work with the co-PI. They study the
problem of penalized empirical risk minimization over the linear
span of N (which can be a large number) of Reproducing Kernel Hilbert
Spaces (RKHS), the complexity penalty being, up to a regularization parameter, the sum of
the p-th powers of the RKHS-norms. The examples include regression and classification
problems with specified groups of covariates (features), aggregation problems in Kernel
Machines learning and high dimensional additive models in regression and classification. It
happens that in this more general situation it is also possible to prove that l_p-type
complexity penalization with p=1 (or close to 1) provides a way to recover a sparse solution of
the problem if it exists.
In the first paper in this direction that appeared in COLT 2008, the PI and the co-PI
studied
the properties of sparse recovery with a large number of reproducing kernel Hilbert
spaces. The method is based on l_1 type
penalization: up to a regularization parameter, the penalty is the sum of RKHS norms of
the components.
This is an infinite dimensional version of LASSO and it is known
to be equivalent to so called multiple kernel learning, a method
developed and studied in Machine Learning literature. The PI and
the co-PI proved oracle inequalities showing that this approach
is adaptive to the underlying sparsity of the problem.

The PI and the co-PI have also studied (in a paper published recently in the Annals of Statistics) another approach
to the problem of multiple kernel learning that is based on penalized empirical
risk minimization with complexity penalty determined jointly by the empirical L_2
norms
and the reproducing kernel Hilbert space (RKHS) norms induced by the kernels and with a
data-driven choice of regularization parameters.
The main focus is on the
case when the total number of kernels is large,
but only a relatively small
number of them is needed to represent the target function, so the problem is
sparse.
They established oracle inequalities for the excess
risk of the resulting prediction rule
that show that the method is
adaptive to the unknown design distribution, to the smoothness
of the components and to the sparsity of the problem.

3. In a joint work with Evarist Gine, the PI investigates the penalized empirical risk
minimization over a RKHS with complexity penalty being the square of the RKHS-norm
(times a regularization parameter). They use the approach by the PI developed in his paper
(in 'Annales I.H.P.') and derive new excess risk bounds as well as new bounds on the
accuracy of solution measured by the RKHS-norm. The bounds depend on the behavior of
the eigenvalues of the operator generated by the kernel and they improve recent results by
Smale and Zhou in the same context.

4. The PI has studied a method of sparse recovery in a convex hull of a given finite
dictionary
based on penalization with negative entropy
of vectors of coefficients. In this context, it was possible to prove
several inequalities showing that sparsity of a penalized risk minimization problem leads
to approximate sparsity of its empirical solution, and, moreover, the L_2-error of the

empirical solution
can be controlled in terms of sparsity of the problem under rather
weak assumptions on the dictionary. In particular, these results
can be applied to a variety of problems of sparse mixture recovery.
The paper on this subject has appeared in the Annals of Statistics.

5. The PI has also proved new oracle inequalities for ℓ_1-penalized empirical risk minimization in terms of several measures of 'alignment' between the dictionary and the vectors of gradient of the penalty or of penalized risk. This is an attempt to develop more subtle ways to describe rather complicated geometric relationships in sparse recovery problems. This approach has been systematically used in the lectures the PI gave at Saint Flour Summer School in July, 2008 and it was developed in the lecture notes that are to be published in Springer Lecture Notes in Mathematics Series.

6. The PI studied sequential algorithms of active learning based on the estimation of the level sets of the empirical risk. Localized
Rademacher complexities are used in the algorithms to estimate
the sample sizes
needed to achieve the required accuracy of learning in an adaptive way.
Probabilistic
bounds on the number of active examples have been proved and several
applications to binary classification problems are considered.
These results have been published in the Journal of Machine
Learning Research.

7. In a joint paper with Yi Lin, the co-PI investigates the consistency and algorithmic aspects
of the nonnegative garrote. It has been shown that the nonnegative garrote is consistent in variable selection whereas several popular alternatives are not. It also has been shown that the nonnegative garrote has a piecewise linear solution path and can therefore be computed efficiently. The paper has appeared in JRSS B earlier this year.

8. In another joint paper with Renato Menteiro, Ali Ekici and Zhaosong Lu, the co-PI proposes a new notion of sparsity for multivariate linear regression where multiple responses are considered simultaneously. The sparsity is closely related to the rank of the coefficient matrix. Lower ranks imply smaller numer of common factors. They propose efficient algorithm to compute the solution. This paper also appeared in JRSS B recently.

9. In a paper appeared in Technometrics, the co-PI, together with Roshan Joseph and Yi Lin propose a modified LARS algorithm that can effectively enforce heredity principles in variable selection. The approach is fast to compute. In a followup work, the co-PI, and Roshan Joseph and Hui Zou, consider more general model building constraints in variable selection, which they refer to as the structured variable selection and estimation. They propose new method for this purpose and show that the proposed estimate is consistent for both variable selection and estimation. The paper has been accepted for publication by the Annals of Applied Statistics.

10. In joint work with Marten Wegkamp, the co-PI studied the problem of classification with
rejection options. In many practical situations, the loss of misclassification may be too heavy and not making a definite classification may be more desirable. We consider mathematical formulation of these problems in terms of convex risk minimization. One of the joint papers has been accepted by Journal of Machine Learning Research subject to minor revision. In another joint work, we study the sparsity problem associated with this type of methods and the manuscript has also been submitted recently.

11. Together with Tony Cai, the co-PI studied the problem of functional linear regression. A
reproducing kernel Hilbert space based approach is proposed as an alternative to the usual functional PCA based approach. As such, the proposed method can avoid some unnecessary condition required in order to recover the functional PCA rather than the slope function itself. It identifies minimal conditions that are required for optimal estimation of the functional linear regression.

12. In joint work with Hui Zou, a new computational approach is proposed to compute a general nonlinear solution path associated with the l1 type regularization. The approach is based on global approximation to the whole solution path and can be employed in several more sophisticated settings. The paper appeared in JASA in 2009.

Training and Development:

Philippe Rigollet was a post doc of the PI in 2007-2008. He has been actively involved in a joint work with the PI and in the research seminar on Sparsity in High Dimensional Problems in Learning Theory
organized by the PI and the co-PI. He has accepted a tenure-track position at Princeton.

Stas Minsker, a PhD student of the PI, was supported in part by this NSF grant and was working under the PI's supervision on problems of sparse recovery in infinite dictionaries. He has successfully passed PhD comprehensive exams at Georgia Tech and is now doing his PhD thesis research.

Lingyan Ruan, a PhD student of the co-PI, was supported in part by this NSF grant and is working under the co-PI's supervision on problems of high dimensional nonparametric modeling and estimation. She has successfully passed PhD comprehensive exams at Georgia Tech and is now doing her PhD thesis research.

Kai Ni, a PhD student of the PI (co-advised with Xiaoming Huo, ISYE, Georgia Tech) is working on the development of excess risk bounds in multiclass problems and on hypotheses testing for filaments
in point processes. He was supported in part by this NSF grant.
He has also successfully passed PhD comprehensive exams and is making progress on his PhD research.

Pedro Rangel, a PhD student of the PI, is currently working
on the development of multiple kernel learning methodology.
He has successfully passed PhD comprehensive exams at Georgia Tech
and he is starting his PhD research.

Outreach Activities:

The PI and the co-PI have been involved in organizing several conferences, workshops, etc related to this project:

1. The PI and the co-PI jointly with Justin Romberg and Xiaoming Huo have organized a Workshop on Sparsity in High Dimensional Statistics
and Learning Theory.

2. The PI served on the Organizing Committee of the Workshop
on Approximation and Learning in High Dimensions at Texas A&M in October, 2007.

3. The PI served on the Organizing Committee of the 5th International Conference on High

Dimensional Probability in C.I.R.M, Luminy, France in May, 2008. It was sponsored by NSF, NSA and C.I.R.M.

4. The co-PI has been serving on the Organizing Committee of AISTAT 2007 held in Puerto Rico in March, 2007.

5. The PI served on the Program Committees of COLT conferences in Helsinki, Finland, 2008 and Haifa, Israel, 2010.

6. The PI was a co-organizer (with Bartlett, Mammen, Tsybakov and van de Geer) of a workshop 'Sparse Recovery Problems in High Dimensions: Statistical Inference and Learning Theory' in Oberwolfach, Germany, 2009.

7. The PI was a co-organizer of a Workshop on Geometric Aspects of Machine Learning and Visual Analytics during IEEE VisWeek in Atlantic City in October, 2009 (supported by another NSF award).

8. The PI is serving on the Scientific Committee of an SMF Workshop on Learning Theory to take place at Poincare Institute in Paris in May, 2011.

The PI has been also serving on the editorial boards of the Annals of Statistics, Bernoulli and Statistics Surveys handling the papers in the research areas closely related to this project. The PI has reviewed numerous papers for these and other journals, reviewed proposals for NSF, NSA and funding agencies in other countries, served on NSF panels.

In the past years, the co-PI has reviewed papers for various statistics journals. The co-PI is also serving on the editorial boards of Biometrics and Annals of Statistics.

Journal Publications

Koltchinskii, V., "Sparsity in Penalized Empirical Risk Minimization", Annales de'l Institut H. Poincare (B), Probabilites et Statistiques, p. 7, vol. 45, (2009). Published,

Koltchinskii, V., "The Dantzig Selector and Sparsity Oracle Inequalities", Bernoulli, p. 799, vol. 15, (2009). Published,

Yuan, M.; Lin, Y., "On the nonnegative garrote estimator", Journal of the royal statistical society, series B, p. 143, vol. 69(2), (2007). Published,

Yuan, M.; Eciki, A.; Lu, Z.; Monteiro, R., "Dimension reduction and coefficient estimation in multivariate linear regression", Journal of the royal statistical society, series B, p. 329, vol. 69(3), (2007). Published,

Yuan, M.; Lin, Y., "Model selection and estimation in the Gaussian graphical model", Biometrika, p. 19, vol. 94, (2007). Published,

Yuan, M.; Roshan, J.; Lin, Y., "An efficient variable selection approach for analyzing designed experiments", Technometrics, p. , vol. , (2007). Accepted,

Yuan, M.; Roshan, J.; Zou, H., "Structured variable selection and estimation", Annals of Applied Statistics, p. 17, vol. 3, (2009). Published,

Zou, H. and Yuan, M., "Regularized Simultaneous Model Selection in Multiple Quantile Regression", Computational Statistics and Data Analysis, p. 529, vol. 52, (2008). Published,

Zou, H. and Yuan, M., "Composite Quantile Regression and the Oracle Model Selection Theory", Annals of Statistics, p. , vol. , (2008). Published,

Yuan, M., "Efficient Computation of the ℓ_1 Regularized Solution Path in Gaussian Graphical Models", Journal of Computational and Graphical Statistics, p. 809, vol. 17, (2008). Published,

Li, J., Yuan, M. and Lee, C., "Approximate Test Risk Bound Minimization through Soft Margin Estimation", IEEE Transactions on Acoustics, Speech, and Signal Processing, p. , vol. 15, (2007). Published,

Wu, S. Zou, H, and Yuan, M., "Structured Variable Selection in Support Vector Machines", Electronic Journal of Statistics, p. , vol. 2, (2008). Published,

Koltchinskii, V., "Sparse Recovery in Convex Hulls via Entropy Penalization", Annals of Statistics, p. 1332, vol. 37, (2009). Published,

Koltchinskii, V., "Rademacher Complexities and Bounding the Excess Risk in Active Learning", Journal of Machine Learning Research, p. 2457, vol. 11, (2010). Published,

Koltchinskii, V. and Yuan, M., "Sparsity in Multiple Kernel Learning", Annals of Statistics, p. 3660, vol. 38, (2010). Published,

Yuan, M. and Wegkamp, M., "Classification Methods with Reject Option Based on Convex Risk Minimization", Journal of Machine Learning Research, p. , vol. 11, (2010). Published,

Yuan, M. and Cai, T. T., "A Reproducing Kernel Hilbert Space Approach to Functional Linear Regression", Annals of Statistics, p. , vol. , (2008). Accepted,

Yuan, M. and Zou, H., "Efficient Global Approximation of Generalized Nonlinear l1 Regularized Solution Paths and Its Applications", Journal of the American Statistical Association, p. , vol. 104, (2009). Published,

Yuan, M., "State Price Density Estimation via Nonparametric Mixtures", Annals of Applied Statistics, p. , vol. 3, (2009). Published,

Yuan, M. and Huang, J., "Regularized Parameter Estimation of High Dimensional t Distribution", Journal of Statistical Planning and Inference, p. 2284, vol. 139, (2009). Published,

Books or Other One-time Publications

Koltchinskii, V, "Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems", (2008). Book, In preparation for submission to Springer
Collection: Lectures in 38th Probability Summer School in Saint-Flour,
2008
Bibliography: Lecture Notes in Mathematics, Springer

Koltchinskii, V. and Sakhanenko, L., "Asymptotics of Statistical Estimators of Integral Curves", (2009). Book, Published
Editor(s): C. Houdre, V. Koltchinskii,
D. Mason, M. Peligrad (Eds)
Collection: High Dimensional Probability, The Luminy Volume, pp. 326--377
Bibliography: IMS Collections

Koltchinskii, V. and Yuan, M., "Sparse recovery in large ensembles of Kernel Machines", (2008). Book, Published
Collection: 21st Annual Conference
on Learning Theory, COLT 2008,
pp. 229--238
Bibliography: Omnipress

Koltchinskii, V. and Minsker, S., "Sparse Recovery in Convex Hulls
of Infinite Dictionaries", (2010). Book, Published

Editor(s): A. Kalai and M. Mohri
Collection: Proceedings of 23rd Conference
on Learning Theory, COLT 2010,
pp. 420--432
Bibliography: Haifa, Israel

Web/Internet Site

Other Specific Products

Contributions

Contributions within Discipline:

This project has contributed to a better understanding of the role of
complexity penalties in sparse recovery problems. Specifically, it has been shown that l_p-type penalties with a proper choice of p provide some adaptation to unknown degree of sparsity of the problem. This holds in the case of large finite dictionaries (even when the dictionary is not 'well posed') as well as in much broader frameworks of high dimensional problems of learning theory. The same applies to some other types of penalties, such as entropy based penalty in the case of sparse recovery in convex hulls. This opened a possibility to study sparse recovery problems for infinite dictionaries that are frequently used in learning theory. In particular, it became possible to develop sparse recovery methods in large ensembles of kernel machines (that include many important models in statistics such as additive models) and for linear spans and convex hulls of
infinite dictionaries (which provides a way to understand sparse mixtures recovery).

Contributions to Other Disciplines:

Based on the results of this project, it becomes possible to develop
much more general methods of sparse recovery that can be applied to
more complex statistical models of high-dimensional data with potential impact in a variety of fields (including image processing, bioinformatics and financial engineering).

Contributions to Human Resource Development:

Contributions to Resources for Research and Education:

Contributions Beyond Science and Engineering:

Conference Proceedings

Categories for which nothing is reported:

Organizational Partners
Any Web/Internet Site
Any Product
Contributions: To Any Human Resource Development
Contributions: To Any Resources for Research and Education
Contributions: To Any Beyond Science and Engineering
Any Conference